



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 9/12, 5/00, 1/20, 15/00, C12P 21/06, 19/34, C07K 1/00, C07H 21/02, 21/04	A1	(11) International Publication Number: WO 98/01542 (43) International Publication Date: 15 January 1998 (15.01.98)
(21) International Application Number: PCT/US97/12296 (22) International Filing Date: 8 July 1997 (08.07.97) (30) Priority Data: 08/676,974 8 July 1996 (08.07.96) US (71) Applicant: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA [US/US]; 22nd floor, 300 Lakeside Drive, Oakland, CA 94612-3550 (US). (72) Inventor: COLLINS, Kathleen; University of California, Berkeley, 401 Barker Hall, Berkeley, CA 94720-3204 (US). (74) Agent: OSMAN, Richard, Aron; Science & Technology Law Group, Suite 3200, 268 Bush Street, San Francisco, CA 94104 (US).		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG). Published <i>With international search report.</i>
(54) Title: HUMAN TELOMERASE		
(57) Abstract <p>The invention provides methods and compositions relating to a human telomerase and related nucleic acids, including four distinct human telomerase subunit proteins called p140, p105, p48 and p43 having human telomerase-specific activity. The proteins may be produced recombinantly from transformed host cells from the disclosed telomerase encoding nucleic acids or purified from human cells. Also included are human telomerase RNA components, as well as specific, functional derivatives thereof. The invention provides isolated telomerase hybridization probes and primers capable of specifically hybridizing with the disclosed telomerase gene, telomerase-specific binding agents such as specific antibodies, and methods of making and using the subject compositions in diagnosis, therapy and in the biopharmaceutical industry.</p>		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakistan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

Human Telomerase

INTRODUCTION

Field of the Invention

The field of this invention is an enzyme involved in cell replication.

5

Background

DNA at chromosome ends is maintained in a dynamic balance of loss and addition of telomeric simple sequence repeats. Sequence loss occurs during cell replication, in part from incomplete replication of chromosome termini by DNA-dependent DNA polymerase. Telomeric repeat addition is catalyzed by the enzyme telomerase: a ribonucleoprotein enzyme which uses a short region within the RNA as a template for the polymerase reaction. Although cells can maintain a constant number of telomeric repeats by balancing repeat loss and addition, not all cells do so. Human germline and cancer cells maintain a constant number of telomeric repeats, while normal human somatic cells lose telomeric repeats with each cycle of cell division. Cells which do not maintain stable telomere length demonstrate a limited proliferative capacity: these cells senesce after a number of population doublings correlated with the erosion of telomeres to a critical minimum length.

10

15

20

Because normal somatic cells do not appear to express or require telomerase and do not maintain chromosome ends, and because all or almost all cancer cells express high levels of telomerase activity and maintain chromosome ends, molecules that inhibit or alter telomerase activity could provide effective and non-toxic anti-cancer agents. Similarly, inhibition of telomerase in parasitic or infectious agents (e.g. trypanosomes, fungi, etc.) could provide a specific approach for reducing the viability or proliferation of these agents.

Conversely, activation of telomerase in proliferation-restricted cells (such as normal somatic cells of the blood, vasculature, liver, skin, etc.) could provide a mechanism for promoting additional proliferative lifespan.

Relevant Literature

Purification of telomerase from the ciliate *Tetrahymena* and cloning of genes encoding two protein components of the enzyme is reported in Collins et al. (1995) *Cell* 81, 677-686 and copending US patent application No. 08/359,125, filed 19 DEC 1994. Literature relating to human telomerase include; Kim et al. (1994) *Science* 266, 2011-2014; and Feng et al. (1995) *Science* 269, 1236-1241. Literature relating to telomerase template modifications include Autexier et al. (1994) *Genes and Devel* 8, 563-575; Yu et al. (1991) *Cell* 67, 823-832; and Yu et al. (1990) *Nature* 344, 126-132. The Washington University-Merck EST Project contains an EST, reportedly deposited by Hillier et al. on Nov 1, 1995, which has sequence similarity with the 3' end of SEQ ID NO:3, disclosed herein. For a general review, see Blackburn et al., Eds. (1995) *Telomeres*, Cold Spring Harbor Laboratory Press.

SUMMARY OF THE INVENTION

The invention provides methods and compositions relating to a human telomerase and related nucleic acids. Included are four distinct human telomerase subunit proteins, called p140, p105, p48 and p43 and telomerase protein domains thereof having telomerase-specific activity. The proteins may be produced recombinantly from transformed host cells from the subject telomerase encoding nucleic acids or purified from human cells. Also included are human telomerase RNA components, as well as specific, functional derivatives thereof.

The invention provides isolated telomerase hybridization probes and primers capable of specifically hybridizing with the disclosed telomerase gene, telomerase-specific binding agents such as specific antibodies, and methods of making and using the subject compositions in diagnosis (e.g. genetic hybridization screens for telomerase transcripts), therapy (e.g. gene therapy to modulate telomerase gene expression) and in the biopharmaceutical industry (e.g. reagents for screening chemical libraries for lead pharmacological agents and nucleic acid polymerase reagents).

SEQ ID LISTING

- SEQ ID NO:1: p105 protein (amino acid sequence)
SEQ ID NO:2: p105 ambiguity maximized synthetic DNA
SEQ ID NO:3: p105 natural cDNA (the coding region is bp 97-2370)
SEQ ID NO:4: p105 E. coli optimized synthetic DNA
5 SEQ ID NO:5: p105 mammalian optimized synthetic DNA
SEQ ID NO:6: telomerase RNA
SEQ ID NO:7: telomerase RNA template region modification 1
SEQ ID NO:8: telomerase RNA template region modification 2
SEQ ID NO:9: telomerase RNA template region modification 3
10 SEQ ID NO:10 p43 peptide (XXXEAAT[I/L]D[I/L]PQQGANK, where the three X's are
indeterminant residues)

DETAILED DESCRIPTION OF THE INVENTION

The invention provides isolated human telomerase proteins including human
15 telomerase proteins p140, p105, p48 and p43, having molecular weights of about 140kD,
about 105kD, about 48kD and about 43kD, respectively, as determined by polyacrylamide
gel electrophoresis under denaturing conditions (Matsudaira and Burgess (1978) Anal
Biochem 87, 386-396), and telomerase protein domains thereof. The telomerase proteins
comprise assay-discernable functional domains including RNA recognition motifs and
20 subunit binding domains and may be provided as fusion products, e.g. with non-telomerase
polypeptides. The human telomerase proteins of the invention, including the subject protein
domains, all have telomerase-specific activity or function.

Telomerase-specific activity or function may be determined by convenient *in vitro*,
cell-based, or *in vivo* assays: e.g. *in vitro* binding assays, cell culture assays, in animals (e.g.
25 immune response, gene therapy, transgenics, etc.), etc. Binding assays encompass any
assay where the molecular interaction of a telomerase protein with a binding target is
evaluated. The binding target may be a natural intracellular binding target such as a
telomerase subunit (e.g. another protein subunit or RNA subunit), a substrate, agonist,
antagonist, chaperone, or other regulator that directly modulates telomerase activity or its
30 localization; or non-natural binding target such as a specific immune protein such as an
antibody, or a telomerase specific agent such as those identified in assays described below.

Generally, telomerase-binding specificity is assayed by telomere polymerase activity (see, e.g. Collins et al. 1995, Cell 81, 677-686), by binding equilibrium constants (usually at least about 10^7 M^{-1} , preferably at least about 10^8 M^{-1} , more preferably at least about 10^9 M^{-1}), by the ability of the subject protein to function as negative mutants in telomerase-expressing cells, to elicit telomerase specific antibody in a heterologous host (e.g. a rodent or rabbit), etc. In any event, the telomerase binding specificity of the subject telomerase proteins necessarily distinguishes ciliate telomerase, preferably distinguishes non-mammalian telomerases and more preferably distinguishes non-human telomerases. Exemplary telomerase proteins which are shown to have telomerase binding specificity include the telomerase RNA (e.g. SEQ ID NO:6) binding domains (e.g. RRM 1-4: SEQ ID NO:1, about residues 5-81, residues 115-192, residues 336-420, and residues 487-578, respectively), telomerase primer binding domains, nucleotide triphosphate binding domains and binding domains of regulators of telomerase such as nuclear localization proteins, etc. As used herein, a protein domain comprises at least 12, preferably at least about 20, more preferably at least about 40, most preferably at least about 80 residues of the disclosed respective SEQ ID NO.

The claimed human telomerase proteins are isolated or pure: an "isolated" protein is unaccompanied by at least some of the material with which it is associated in its natural state, preferably constituting at least about 0.5%, and more preferably at least about 5% by weight of the total protein in a given sample and a pure protein constitutes at least about 90%, and preferably at least about 99% by weight of the total protein in a given sample. The telomerase proteins and protein domains may be synthesized, produced by recombinant technology, or purified from human cells. A wide variety of molecular and biochemical methods are available for biochemical synthesis, molecular expression and purification of the subject compositions, see e.g. Molecular Cloning, A Laboratory Manual (Sambrook, *et al.* Cold Spring Harbor Laboratory), Current Protocols in Molecular Biology (Eds. Ausubel, *et al.*, Greene Publ. Assoc., Wiley-Interscience, NY) or that are otherwise known in the art. An exemplary method for isolating each of human telomerase protein p140, p105, p48 and p43 from human cells is as follows:

Several thousand (two to twelve thousand) liters of HeLa cells are grown in spinner culture. The cells are lysed by dounce homogenization in low-salt buffer to produce crude cell lysates. The lysates are supplemented with 15% glycerol and centrifuged at 125,000 x

g for 50 minutes to obtain a first soluble fraction enriched for telomerase activity (S-100 fraction). The S-100 fraction is adjusted to 0.2 M ammonium sulfate, bound to SP Sepharose (Pharmacia), and developed with a gradient in sodium chloride, to obtain a second soluble fraction enriched for telomerase (SP fraction). The SP fraction is adjusted to about 0.3-0.4 M ionic strength and bound to Q Sepharose (Pharmacia), and developed with a gradient in sodium chloride, to obtain a third soluble fraction enriched for telomerase (Q fraction). The Q fraction is adjusted to about 0.3-0.4 M ionic strength, bound to phosphocellulose (Whatman), and developed with sodium chloride, to obtain a fourth soluble fraction enriched for telomerase (PC fraction). The PC fraction is adjusted to about 0.3-0.4 M ionic strength, bound to 2'Omethyl RNA oligonucleotide immobilized on streptavidin agarose (Sigma), and eluted with a electrophoresis sample medium comprising 5% β -mercaptoethanol and 2% Sodium Dodecyl Sulfate to obtain a fifth soluble fraction (2'Omethyl fraction). The 2'Omethyl fraction is separated by polyacrylamide gel electrophoresis under denaturing conditions (Matsudaira and Burgess (1978) Anal Biochem 87, 386-396) to obtain gel protein bands at a molecular weight of about 140kD, 105kD, 48kD or 43kD having telomerase activity. The gel bands are excised or blotted to obtain purified human telomerase proteins p140, p105, p48 and p43.

The subject telomerase proteins find a wide variety of uses including use in isolating, enriching for and concentrating telomerase RNA and telomerase proteins, as immunogens, in the methods and applications described below, as reagents in the biotechnology industries, and in therapy. Recombinant telomerases are used in many applications where nascent oligonucleotides of predetermined sequence are desired. For example, native nucleic acid molecules are labeled or extended at their 3' ends by addition of a predetermined repeat sequence (for double-stranded oligonucleotides, both ends of the molecule may be tagged). Oligonucleotides complementary to the repeat are then used to amplify, sequence, affinity purify, etc. the nucleic acid molecules. The use of a repeat sequence for 3' end tagging improves specificity and provides sequence alternatives compared with non-templated enzymes presently available for this purpose, e.g. terminal transferase. Repeats encoding restriction enzyme sites provide repeat tagging to facilitate cloning and the use of telomerase alleviates the restrictive conditions required for optimal ligation with available ligase enzymes. Telomerase also finds use in regulating cell growth

or increasing cell density tolerance; for example, cells contacted with an effective amount of exogenous telomerase to overcome the growth control limitation otherwise imposed by short telomere length. Telomerase may be introduced, expressed, or repressed in specific populations of cells by any convenient way such as microinjection, promoter-specific expression of recombinant enzyme, targeted delivery of lipid vesicles, etc. Advantageously, only a brief period of telomerase activity is required to allow many generations of continued proliferation of the contacted cell, due to the ability of telomerase to extend telomeres in one cell cycle by more sequence than is lost with each cell division.

The invention provides natural and non-natural human telomerase-specific binding agents including substrates, agonist, antagonist, etc., methods of identifying and making such agents, and their use in diagnosis, therapy and pharmaceutical development. For example, human telomerase-specific agents are useful in a variety of diagnostic and therapeutic applications. Novel human telomerase-specific binding agents include human telomerase-specific receptors, such as somatically recombined protein receptors like specific antibodies or T-cell antigen receptors (see, e.g. Harlow and Lane (1988) Antibodies, A Laboratory Manual, Cold Spring Harbor Laboratory) and other natural intracellular binding agents identified with assays such as one-, two- and three-hybrid screens, non-natural intracellular binding agents identified in screens of chemical libraries such as described below, etc. Agents of particular interest modulate human telomerase function, e.g. human telomerase antagonists and find use methods for modulating the binding of a human telomerase or telomerase protein to a human telomerase binding target.

For diagnostic uses, the binding agents are frequently labeled, such as with fluorescent, radioactive, chemiluminescent, or other easily detectable molecules, either conjugated directly to the binding agent or conjugated to a probe specific for the binding agent. Binding agents also find use in modulating the telomerase activity present in a cell. For example, isolated cells, whole tissues, or individuals may be treated with a telomerase binding agent to activate, inhibit, or alter the specificity of telomerase assembly, localization, substrate interaction, or synthesis activity. Effectively treated cells have increased or decreased replication potential, or suffer from loss of proper telomere structure (resulting in lethality). These binding agents also find therapeutic use to control cell proliferation; for example, the uncontrolled growth of transformed cells (e.g. cancer cells) is managed by administration to the cells or patient comprising such cells of a telomerase

binding agent which reduces telomerase activity. In contrast to many current chemotherapies, the present invention provides enhanced specificity of lethality, with minimum toxicity to dividing yet normal somatic cells.

The amino acid sequences of the disclosed telomerase proteins are used to back-translate telomerase protein-encoding nucleic acids optimized for selected expression systems (Holler et al. (1993) Gene 136, 323-328; Martin et al. (1995) Gene 154, 150-166) or used to generate degenerate oligonucleotide primers and probes for use in the isolation of natural telomerase encoding nucleic acid sequences ("GCG" software, Genetics Computer Group, Inc, Madison WI). As examples, SEQ ID NO:2 discloses an ambiguity-maximized p105 coding sequence encompassing all possible nucleic acids encoding the full-length p105 protein. SEQ ID NO:3 discloses a natural human cDNA sequence encoding p105. SEQ ID NO:4 is a p105 coding sequence codon-optimized for E. coli, SEQ ID NO:5 is a p105 coding sequence codon optimized for mammalian cell expression. Telomerase encoding nucleic acids may be part of human telomerase-expression vectors and may be incorporated into recombinant host cells, e.g. for expression and screening, transgenic animals, e.g. for functional studies such as the efficacy of candidate drugs for disease associated with human telomerase-mediated signal transduction, etc. Expression systems are selected and/or tailored to effect human telomerase protein structural and functional variants through alternative post-translational processing.

The invention also provides nucleic acid hybridization probes and replication/amplification primers having a human telomerase cDNA specific sequence contained in SEQ ID NO:3, bases 1-2345, and sufficient to effect specific hybridization thereto (i.e. specifically hybridize with SEQ ID NO:3, bases 1-2345 in the presence of natural ciliate telomerase cDNA, preferably in the presence of non-mammalian telomerase cDNA and more preferably, in the presence of murine telomerase cDNA). Demonstrating specific hybridization generally requires stringent conditions, for example, hybridizing in a buffer comprising 30% formamide in 5 x SSPE (0.18 M NaCl, 0.01 M NaPO₄, pH7.7, 0.001 M EDTA) buffer at a temperature of 42°C and remaining bound when subject to washing at 42°C with 0.2 x SSPE; preferably hybridizing in a buffer comprising 50% formamide in 5 x SSPE buffer at a temperature of 42°C and remaining bound when subject to washing at 42°C with 0.2 x SSPE buffer at 42°C. Human telomerase cDNA homologs can also be distinguished from other protein using alignment algorithms, such as BLASTX (Altschul *et*

al. (1990) Basic Local Alignment Search Tool, *J Mol Biol* 215, 403-410).

The invention also provides non-natural sequence, recombinant and isolated natural sequence human telomerase RNA. Natural human telomerase RNA sequences include the nucleic acid disclosed as SEQ ID NO:6, or a fragment thereof sufficient to specifically hybridize with a nucleic acid having the sequence defined by SEQ ID NO:6 in the presence of a nucleic acid having the sequence disclosed in Feng et al. 1995, *Science* 269, 1236-1241. Such fragments necessarily distinguish the previously described (Feng et al. 1995, *Science* 269, 1236-1241) human RNA species. Preferred such fragments comprise SEQ ID NO:6, bases 191-210, bases 245-259, bases 341-369 or bases 381-399. Non-natural sequences include derivatives and/or mutations of SEQ ID NO:6, where such derivatives/mutations provide alteration in template, protein binding, or other regions to effect altered telomerase substrate specificity or altered reaction product (e.g. any predetermined sequence), etc.; see, e.g. Autexier et al., 1994, *Genes & Develop* 8, 563-575; Collins et al. (1995) *EMBO J.* 14, 5422-5432; Greider et al. (1995) *Structure and Biochemistry of Ciliate and Mammalian Telomerases*, in *DNA Replication*, DePamphilis, Ed., Cold Spring Harbor Laboratory Press. Additional derivatives function as dominant negative fragments which effectively compete for telomerase assembly. For examples, SEQ ID NO:7, 8 and 9 are derivatives which provide for modified substrate specificity and polymerase reaction product to interfere with cellular function (see, e.g. Hanish et al. (1994) *Proc Natl Acad Sci* 91, 8861-8865).

The subject nucleic acids are of synthetic/non-natural sequences and/or are isolated, i.e. unaccompanied by at least some of the material with which it is associated in its natural state, preferably constituting at least about 0.5%, preferably at least about 5% by weight of total nucleic acid present in a given fraction, and usually recombinant, meaning they comprise a non-natural sequence or a natural sequence joined to nucleotide(s) other than that which it is joined to on a natural chromosome. Nucleic acids comprising the nucleotide sequence of SEQ ID NO:3 or fragments thereof, contain such sequence or fragment at a terminus, immediately flanked by a sequence other than that which it is joined to on a natural chromosome, or flanked by a native flanking region fewer than 10 kb, preferably fewer than 2 kb, which is immediately flanked by a sequence other than that which it is joined to on a natural chromosome. While the nucleic acids are usually RNA or DNA, it is often advantageous to use nucleic acids comprising other bases or nucleotide analogs to

provide modified stability, etc. The subject nucleic acids find a wide variety of applications including use as translatable transcripts, hybridization probes, PCR primers, diagnostic nucleic acids, etc.; use in detecting the presence of human telomerase genes and gene transcripts and in detecting or amplifying nucleic acids encoding additional human telomerase homologs and structural analogs.

5 In diagnosis, human telomerase hybridization probes find use in identifying wild-type and mutant human telomerase alleles in clinical and laboratory samples. Mutant alleles are used to generate allele-specific oligonucleotide (ASO) probes for high-throughput clinical diagnoses. In therapy, therapeutic human telomerase nucleic acids are used to modulate cellular expression or intracellular concentration or availability of active
10 telomerase. A wide variety of indications may be treated, either prophylactically or therapeutically with the subject compositions. For example, where limitation of cell growth is desired, e.g. neoproliferative disease, a reduction in telomerase expression is effected by introducing into the targeted cell type human telomerase nucleic acids which reduce the functional expression of human telomerase gene products (e.g. nucleic acids capable of
15 inhibiting translation of a functional telomerase transcript). Conditions for treatment include various cancers, where any of a wide variety of cell types may be involved, restenosis, where vascular smooth muscle cells are involved, inflammatory disease states, where endothelial cells, inflammatory cells and glomerular cells are involved, myocardial infarction, where heart muscle cells are involved, glomerular nephritis, where kidney cells
20 are involved, transplant rejection where endothelial cells are involved, infectious diseases such as HIV infection where certain immune cells and other infected cells are involved, or the like.

 Telomerase inhibitory nucleic acids are typically antisense: single-stranded sequences comprising complements of the disclosed natural telomerase coding sequences.
25 Antisense modulation of the expression of a given telomerase protein may employ telomerase antisense nucleic acids operably linked to gene regulatory sequences. Cells are transfected with a vector comprising a human telomerase sequence with a promoter sequence oriented such that transcription of the gene yields an antisense transcript capable of binding to endogenous human telomerase protein encoding mRNA. Transcription of the
30 antisense nucleic acid may be constitutive or inducible and the vector may provide for stable extrachromosomal maintenance or integration. Alternatively, single-stranded

antisense nucleic acids that bind to genomic DNA or mRNA encoding a given human telomerase protein may be administered to the target cell, in or temporarily isolated from a host, at a concentration that results in a substantial reduction in expression of the targeted protein.

In other indications, e.g. certain hypersensitivities, atrophic diseases, etc., an increase in cell growth or proliferation is desired. In these applications, an enhancement in human telomerase expression is effected by introducing into the targeted cell type human telomerase nucleic acids which increase the functional expression of human telomerase gene products. Conditions for treatment include multiple sclerosis, where certain neuronal cells are involved, inflammatory disease states such as rheumatoid arthritis, where bystander cells are involved, transplant rejection where graft cells are involved, infectious diseases such as HIV infection where certain uninfected host cells are involved, or the like. Such nucleic acids may be human telomerase expression vectors, vectors which upregulate the functional expression of an endogenous human telomerase allele, or replacement vectors for targeted correction of human telomerase mutant alleles.

Various techniques may be employed for introducing of the nucleic acids into viable cells, e.g. transfection with a retrovirus, viral coat protein-liposome mediated transfection. The techniques vary depending upon whether one is using the subject compositions in culture or *in vivo* in a host. In some situations it is desirable to provide the nucleic acid source with an agent which targets the target cells, such as an antibody specific for a surface membrane protein on the target cell, a ligand for a receptor on the target cell, etc. Where liposomes are employed, proteins which bind to a surface membrane protein associated with endocytosis may be used for targeting and/or to facilitate uptake, e.g. capsid proteins or fragments thereof tropic for a particular cell type, antibodies for proteins which undergo internalization in cycling, proteins that target intracellular localization and enhance intracellular half-life.

The invention provides methods and compositions for enhancing the yield of many recombinantly produced proteins by increasing maximum cell densities and survival time of host production cells in culture. Specifically, cultured cells are transfected with nucleic acids which effect the up-regulation of endogenous telomerase or the expression of an exogenous telomerase. For example, nucleic acids encoding functional human telomerase operably linked to a transcriptional promoter are used to over-express the exogenous

telomerase in the host cell. Telomerase-expressing cells demonstrate enhanced survival ability at elevated cell densities and over extended culture periods.

The invention provides efficient methods of identifying agents, compounds or lead compounds for agents active at the level of a human telomerase modulatable cellular function. Generally, these screening methods involve assaying for compounds which modulate human telomerase interaction with a natural human telomerase binding target. A wide variety of assays for binding agents are provided including labeled *in vitro* telomere polymerase assays, protein-protein binding assays, immunoassays, cell based assays, etc. The methods are amenable to automated, cost-effective high throughput screening of chemical libraries for lead compounds. Identified reagents find use in the pharmaceutical industries for animal and human trials; for example, the reagents may be derivatized and rescreened in *in vitro* and *in vivo* assays to optimize activity and minimize toxicity for pharmaceutical development. Target indications may include infection, genetic disease, cell growth and regulatory dysfunction, such as neoplasia, inflammation, hypersensitivity, etc. Target cells also include progenitor cells for repopulating blood or bone marrow, tissue grafts, and tissue subject to degradation/high turnover such as digestive and vascular endothelia and pulmonary and dermal epithelia.

In vitro binding assays employ a mixture of components including a human telomerase protein, which may be part of multi-subunit telomerase, a fusion product with another peptide or polypeptide, e.g. a tag for detection or anchoring, etc. The assay mixtures comprise a natural intracellular human telomerase binding target, e.g. a substrate. While native binding targets may be used, it is frequently preferred to use portions (e.g. peptides, nucleic acid fragments) thereof so long as the portion provides binding affinity and avidity to the subject human telomerase conveniently measurable in the assay. The assay mixture also comprises a candidate pharmacological agent. Candidate agents encompass numerous chemical classes, though typically they are organic compounds; preferably small organic compounds and are obtained from a wide variety of sources including libraries of synthetic or natural compounds. A variety of other reagents may also be included in the mixture. These include reagents like salts, buffers, neutral proteins, e.g. albumin, detergents, protease inhibitors, nuclease inhibitors, antimicrobial agents, etc. may be used.

The resultant mixture is incubated under conditions whereby, but for the presence of

the candidate pharmacological agent, the human telomerase specifically binds the cellular binding target, portion or analog with a reference binding affinity. The mixture components can be added in any order that provides for the requisite bindings and incubations may be performed at any temperature which facilitates optimal binding. Incubation periods are likewise selected for optimal binding but also minimized to facilitate rapid, high-throughput screening.

After incubation, the agent-biased binding between the human telomerase and one or more binding targets is detected by any convenient way. For cell-free binding type assays, a separation step is often used to separate bound from unbound components. Separation may be effected by precipitation (e.g. TCA precipitation, immunoprecipitation, etc.), immobilization (e.g. on a solid substrate), etc., followed by washing by, for examples, membrane filtration (e.g. Whatman's P-81 ion exchange paper, Polyfiltronic's hydrophobic GFC membrane, etc.), gel chromatography (e.g. gel filtration, affinity, etc.). For telomere polymerase assays, binding is detected by a change in the polymerization by the telomerase of a nucleic acid or nucleic acid analog on the substrate.

Detection may be effected in any convenient way. For cell-free binding assays, one of the components usually comprises or is coupled to a label. The label may provide for direct detection as radioactivity, luminescence, optical or electron density, etc. or indirect detection such as an epitope tag, an enzyme, etc. A variety of methods may be used to detect the label depending on the nature of the label and other assay components, e.g. through optical or electron density, radiative emissions, nonradiative energy transfers, etc. or indirectly detected with antibody conjugates, etc.

A difference in the binding affinity of the human telomerase protein to the target in the absence of the agent as compared with the binding affinity in the presence of the agent indicates that the agent modulates the binding of the human telomerase protein to the human telomerase binding target. Analogously, in the cell-based transcription assay also described below, a difference in the human telomerase transcriptional induction in the presence and absence of an agent indicates the agent modulates human telomerase-induced transcription. A difference, as used herein, is statistically significant and preferably represents at least a 50%, more preferably at least a 90% difference.

The following examples are offered by way of illustration and not by way of limitation.

EXAMPLES

1. Protocol for high-throughput human telomere polymerization assay.
- A. Reagents:
 - Neutralite Avidin: 20 µg/ml in PBS.
 - human telomerase: 10^{-8} - 10^{-5} M human telomerase in PBS.
 - 5 - Blocking buffer: 5% BSA, 0.5% Tween 20 in PBS; 1 hour at room temperature.
 - Assay Buffer: 100 mM KCl, 20 mM HEPES pH 7.6, 1 mM MgCl₂, 1 mM dATP, 1 mM dTTP, 1% glycerol, 0.5% NP-40, 50 mM BME, 1 mg/ml BSA, cocktail of protease inhibitors.
 - [³²P]α-dGTP 10x stock: 2×10^{-5} M "cold" dGTP with 100 µCi [³²P]α-dGTP.
 - 10 Place in the 4°C microfridge during screening.
 - telomerase substrate: 10^{-7} - 10^{-4} M biotinylated telomerase substrate (5'-biotin-d(TTAGGG)₃-3') in PBS.
 - Protease inhibitor cocktail (1000X): 10 mg Trypsin Inhibitor (BMB # 109894), 10 mg Aprotinin (BMB # 236624), 25 mg Benzamidine (Sigma # B-6506), 25 mg Leupeptin
 - 15 (BMB # 1017128), 10 mg APMSF (BMB # 917575), and 2mM NaVO₃ (Sigma # S-6508) in 10 ml of PBS.
- B. Preparation of assay plates:
 - Coat with 120 µl of stock N Avidin per well overnight at 4°C.
 - Wash 2 times with 200 µl PBS.
 - 20 - Block with 150 µl of blocking buffer.
 - Wash 2 times with 200 µl PBS.
- C. Assay:
 - Add 40 µl assay buffer/well.
 - Add 40 µl human telomerase (1-1000 fmoles/40 ul in assay buffer)
 - 25 - Add 10 µl compound or extract.
 - Add 10 µl [³²P]α-dGTP 10x stock.
 - Add 40 µl biotinylated telomerase substrate (0.1-10 pmoles/40 ul in assay buffer)
 - Shake at 25°C for 15 minutes.
 - Incubate additional 45 minutes at 25°C.
 - 30 - Stop the reaction by washing 4 times with 200 µl PBS.
 - Add 150 µl scintillation cocktail.

- Count in Topcount.

D. Controls for all assays (located on each plate):

- a. Non-specific binding
- b. cold dGTP at 80% inhibition.

5 2. Protocol for high throughput human telomerase subunit- RNA complex formation assay.

A. Reagents:

- Neutralite Avidin: 20 µg/ml in PBS.
- Blocking buffer: 5% BSA, 0.5% Tween 20 in PBS; 1 hour at room temperature.

10 - Assay Buffer: 100 mM KCl, 20 mM HEPES pH 7.6, 1 mM MgCl₂, 1% glycerol.
0.5% NP-40, 50 mM β-mercaptoethanol, 1 mg/ml BSA, cocktail of protease inhibitors.

- ³²P human telomerase protein 10x stock: 10⁻⁸ - 10⁻⁶ M "cold" human telomerase subunit (p105) supplemented with 200,000-250,000 cpm of labeled human telomerase (Beckman counter). Place in the 4°C microfridge during screening.

15 - Protease inhibitor cocktail (1000X): 10 mg Trypsin Inhibitor (BMB # 109894), 10 mg Aprotinin (BMB # 236624), 25 mg Benzamidine (Sigma # B-6506), 25 mg Leupeptin (BMB # 1017128), 10 mg APMSF (BMB # 917575), and 2mM NaVO₃ (Sigma # S-6508) in 10 ml of PBS.

- telomerase RNA: 10⁻⁷ - 10⁻⁴ M biotinylated RNA (SEQ ID NO:6) in PBS.

20 B. Preparation of assay plates:

- Coat with 120 µl of stock N-Avidin per well overnight at 4°C.
- Wash 2 times with 200 µl PBS.
- Block with 150 µl of blocking buffer.
- Wash 2 times with 200 µl PBS.

25 C. Assay:

- Add 40 µl assay buffer/well.
- Add 10 µl compound or extract.
- Add 10 µl ³³P-human telomerase protein (20,000-25,000 cpm/0.1-10 pmoles/well = 10⁻⁹ - 10⁻⁷ M final concentration).
- Shake at 25°C for 15 minutes.
- Incubate additional 45 minutes at 25°C.

- Add 40 μ l biotinylated RNA (0.1-10 pmoles/40 μ l in assay buffer)
- Incubate 1 hour at room temperature.
- Stop the reaction by washing 4 times with 200 μ l PBS.
- Add 150 μ l scintillation cocktail.
- Count in Topcount.

- 5 D. Controls for all assays (located on each plate):
- a. Non-specific binding
 - b. Soluble (non-biotinylated telomerase) at 80% inhibition.

10 All publications and patent applications cited in this specification are herein
incorporated by reference as if each individual publication or patent application were
specifically and individually indicated to be incorporated by reference. Although the
foregoing invention has been described in some detail by way of illustration and example
for purposes of clarity of understanding, it will be readily apparent to those of ordinary skill
in the art in light of the teachings of this invention that certain changes and modifications
15 may be made thereto without departing from the spirit or scope of the appended claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Kathleen Collins
- (ii) TITLE OF INVENTION: Human Telomerase
- (iii) NUMBER OF SEQUENCES: 10
- 5 (iv) CORRESPONDENCE ADDRESS:
- (A) ADDRESSEE: Science & Technology Law Group
- (B) STREET: 268 Bush Street, Suite 3200
- (C) CITY: San Francisco
- (D) STATE: CA
- 10 (E) COUNTRY: USA
- (F) ZIP: 94104
- (v) COMPUTER READABLE FORM:
- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- 15 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
- (A) APPLICATION NUMBER:
- (B) FILING DATE:
- 20 (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
- (A) NAME: Osman Ph.D., Richard A
- (B) REGISTRATION NUMBER: 36,627
- (C) REFERENCE/DOCKET NUMBER: UCB96-055
- 25 (ix) TELECOMMUNICATION INFORMATION:
- (A) TELEPHONE: (415) 343-4341
- (B) TELEFAX: (415) 343-4342

(2) INFORMATION FOR SEQ ID NO:1:

- 30 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 759 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: not relevant
- 35 (ii) MOLECULE TYPE: peptide
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

	Met	Ala	Gly	Leu	Thr	Leu	Phe	Val	Gly	Arg	Leu	Pro	Pro	Ser	Ala	Arg
	1				5					10					15	
	Ser	Glu	Gln	Leu	Glu	Glu	Leu	Phe	Ser	Gln	Val	Gly	Pro	Val	Lys	Gln
			20						25				30			
	Cys	Phe	Val	Val	Thr	Glu	Lys	Gly	Ser	Lys	Ala	Cys	Arg	Gly	Phe	Gly
5			35					40					45			
	Tyr	Val	Thr	Phe	Ser	Met	Leu	Glu	Asp	Val	Gln	Arg	Ala	Leu	Lys	Glu
		50					55					60				
	Ile	Thr	Thr	Phe	Glu	Gly	Cys	Lys	Ile	Asn	Val	Thr	Val	Ala	Lys	Lys
	65				70					75				80		
10	Lys	Leu	Arg	Asn	Lys	Thr	Lys	Glu	Lys	Gly	Lys	Asn	Glu	Asn	Ser	Glu
				85						90				95		
	Cys	Pro	Lys	Lys	Glu	Pro	Lys	Ala	Lys	Lys	Ala	Lys	Val	Ala	Asp	Lys
			100						105				110			
	Lys	Ala	Arg	Leu	Ile	Ile	Arg	Asn	Leu	Ser	Phe	Lys	Cys	Ser	Glu	Asp
15			115					120					125			
	Asp	Leu	Lys	Thr	Val	Phe	Ala	Gln	Phe	Gly	Ala	Val	Leu	Glu	Val	Asn
		130					135					140				
	Ile	Pro	Arg	Lys	Pro	Asp	Gly	Lys	Met	Arg	Gly	Phe	Gly	Phe	Val	Gln
	145				150					155				160		
20	Phe	Lys	Asn	Leu	Leu	Glu	Ala	Gly	Lys	Ala	Leu	Lys	Gly	Met	Asn	Met
				165						170				175		
	Lys	Glu	Ile	Lys	Gly	Arg	Thr	Val	Ala	Val	Asp	Trp	Ala	Val	Ala	Lys
			180						185				190			
	Asp	Lys	Tyr	Lys	Asp	Thr	Gln	Ser	Val	Ser	Ala	Ile	Gly	Glu	Glu	Lys
25			195						200				205			
	Ser	His	Glu	Ser	Lys	His	Gln	Glu	Ser	Val	Lys	Lys	Lys	Gly	Arg	Glu
		210					215						220			
	Glu	Glu	Asp	Met	Glu	Glu	Glu	Glu	Asn	Asp	Asp	Asp	Asp	Asp	Asp	Asp
	225				230					235				240		
30	Asp	Glu	Glu	Asp	Gly	Val	Phe	Asp	Asp	Glu	Asp	Glu	Glu	Glu	Glu	Asn
				245						250				255		
	Ile	Glu	Ser	Lys	Val	Thr	Lys	Pro	Val	Gln	Ile	Gln	Lys	Arg	Ala	Val
			260						265				270			
	Lys	Arg	Pro	Ala	Pro	Ala	Lys	Ser	Ser	Asp	His	Ser	Glu	Glu	Asp	Ser
35			275						280				285			
	Asp	Leu	Glu	Glu	Ser	Asp	Ser	Ile	Asp	Asp	Gly	Glu	Glu	Leu	Ala	Gln
		290						295					300			

Ser Asp Thr Ser Thr Glu Glu Gln Glu Asp Lys Ala Val Gln Val Ser
 305 310 315 320
 Asn Lys Lys Lys Arg Lys Leu Pro Ser Asp Val Asn Glu Gly Lys Thr
 325 330 335
 Val Phe Ile Arg Asn Leu Ser Phe Asp Ser Glu Glu Glu Glu Leu Gly
 5 340 345 350
 Glu Leu Leu Gln Gln Phe Gly Glu Leu Lys Tyr Val Arg Ile Val Leu
 355 360 365
 His Pro Asp Thr Glu His Ser Lys Gly Cys Ala Phe Ala Gln Phe Met
 370 375 380
 10 Thr Gln Glu Ala Ala Gln Lys Cys Leu Leu Ala Ala Ser Pro Glu Asn
 385 390 395 400
 Glu Ala Gly Gly Leu Lys Leu Asp Gly Arg Gln Leu Lys Val Asp Leu
 405 410 415
 Ala Val Thr Arg Asp Glu Ala Ala Lys Leu Gln Thr Thr Lys Val Lys
 15 420 425 430
 Lys Pro Thr Gly Thr Arg Asn Leu Tyr Leu Ala Arg Glu Gly Leu Ile
 435 440 445
 Arg Ala Gly Thr Lys Ala Ala Glu Gly Val Ser Ala Ala Asp Met Ala
 450 455 460
 20 Lys Arg Glu Arg Phe Glu Leu Leu Lys His Gln Lys Leu Lys Asp Gln
 465 470 475 480
 Asn Ile Phe Val Ser Arg Thr Arg Leu Cys Leu His Asn Leu Pro Lys
 485 490 495
 Ala Val Asp Asp Lys Gln Leu Arg Lys Leu Leu Leu Ser Ala Thr Ser
 25 500 505 510
 Gly Glu Lys Gly Val Arg Ile Lys Glu Cys Arg Val Met Arg Asp Leu
 515 520 525
 Lys Gly Val His Gly Asn Met Lys Gly Gln Ser Leu Gly Tyr Ala Phe
 530 535 540
 30 Ala Glu Phe Gln Glu His Glu His Ala Leu Lys Ala Leu Arg Leu Ile
 545 550 555 560
 Asn Asn Asn Pro Glu Ile Phe Gly Pro Leu Lys Arg Pro Ile Val Glu
 565 570 575
 Phe Ser Leu Glu Asp Arg Arg Lys Leu Lys Met Lys Glu Leu Arg Ile
 35 580 585 590
 Gln Arg Ser Leu Gln Lys Met Arg Ser Lys Pro Ala Thr Gly Glu Pro
 595 600 605

Gln Lys Gly Gln Pro Glu Pro Ala Lys Asp Gln Gln Gln Lys Ala Ala
 610 615 620
 Gln His His Thr Glu Glu Gln Ser Lys Val Pro Pro Glu Gln Lys Arg
 625 630 635 640
 Lys Ala Gly Ser Thr Ser Trp Thr Gly Phe Gln Thr Lys Ala Glu Val
 5 645 650 655
 Glu Gln Val Glu Leu Pro Asp Gly Lys Lys Arg Arg Lys Val Leu Ala
 660 665 670
 Leu Pro Ser His Arg Gly Pro Lys Ile Arg Leu Arg Asp Lys Gly Lys
 675 680 685
 Val Lys Pro Val His Pro Lys Lys Pro Lys Pro Gln Ile Asn Gln Trp
 10 690 695 700
 Lys Gln Glu Lys Gln Gln Leu Ser Ser Glu Gln Val Ser Arg Lys Lys
 705 710 715 720
 Ala Lys Gly Asn Lys Thr Glu Thr Arg Phe Asn Gln Leu Val Glu Gln
 15 725 730 735
 Tyr Lys Gln Lys Leu Leu Gly Pro Ser Lys Gly Ala Pro Leu Ala Lys
 740 745 750
 Arg Ser Lys Trp Phe Asp Ser
 755

20

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2277 base pairs

(B) TYPE: nucleic acid

25 (C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

ATGGCNGGNY TNACNYTNTT YGTNGGNMGN YTNCCNCCNW SNGCNMGNWS NGARCARYTN 60
 30 GARGARYTNT TYWSNCARGT NGGNCCNGTN AARCARTGYT TYGTNGTNAC NGARAARGGN 120
 WSNAARGCNT GYMGNGGNTT YGGNTAYGTN ACNTTYWSNA TGYTNGARGA YGTNCARMGN 180
 GCNYTNAARG ARATHACNAC NTTYGARGGN TGYAARATHA AYGTNACNGT NGCNAARAAR 240
 AARYTNMGNA AYAARACNAA RGARAARGGN AARAAYGARA AYWSNGARTG YCCNAARAAR 300
 GARCCNAARG CNAARAARGC NAARGTNGCN GAYAARAARG CNMGNYTNAT HATHMGNAAY 360
 35 YTNWSNTTYA ARTGYWSNGA RGAYGAYYTN AARACNGTNT TYGCNCARTT YGGNGCNGTN 420
 YTNARGGTNA AYATHCCNMG NAARCCNGAY GGNAARATGM GNGGNTTYGG NTTYGTNCAR 480
 TTYAARAAYY TNYTNGARGC NGGNAARGCN YTNAARGGNA TGAAYATGAA RGARATHAAR 540

	GGNMGNACNG TNGCNGTNGA YTGGGCNGTN GCNAARGAYA ARTAYAARGA YACNCARWSN	600
	GTNWSNGCNA THGGNGARGA RAARWSNCAY GARWSNAARC AYCARGARWS NGTNAARAAR	660
	AARGGNMGNG ARGARGARGA YATGGARGAR GARGARAAYG AYGAYGAYGA YGAYGAYGAY	720
	GAYGARGARG AYGNGTNTT YGAYGAYGAR GAYGARGARG ARGARAAYAT HGARWSNAAR	780
	GTNACNAARC CNGTNCARAT HCARAARMGN GCNGTNAARM GNCCNGCNCC NGCNAARWSN	840
5	WSNGAYCAYW SNGARGARGA YWSNGAYYTN GARGARWSNG AYWSNATHGA YGAYGGNGAR	900
	GARYTNGCNC ARWSNGAYAC NWSNACNGAR GARCARGARG AYAARGCNGT NCARGTNWSN	960
	AAYAARAARA ARMGNAARYT NCCNWSNGAY GTNAAYGARG GNAARACNGT NTTYATHMGN	1020
	AAYYTNSNT TYGAYWSNGA RGARGARGAR YTNGGNGARY TNYTNCARCA RTTYGGNGAR	1080
	YTNAARTAYG TNMGNATHGT NYTNCAYCCN GAYACNGARC AYWSNAARGG NTGYGCNTTY	1140
10	GCNCARTTYA TGACNCARGA RGCNGCNCAR AARTGYTTY TNGCNGCNWS NCCNGARAA	1200
	GARGCNGGNG GNYTNAARYT NGAYGGNMGN CARYTNAARG TNGAYYTNGC NGTNACNMGN	1260
	GAYGARGCNG CNAARYTNCA RACNACNAAR GTNAARAARC CNACNGGNAC NMGNAAYYTN	1320
	TAYYTNGCNM GNGARGGNYT NATHMGNGCN GGNACNAARG CNGCNGARGG NGTNWSNGCN	1380
	GCNGAYATGG CNAARMGNGA RMGNTTYGAR YTNYTNAARC AYCARAARYT NAARGAYCAR	1440
15	AAYATHHTYG TNWSNMGNAC NMGNYTNTGY YTNCAAYAYY TNCCNAARGC NGTNGAYGAY	1500
	AARCARYTNM GNAARYTNYT NYTNWSNGCN ACNWSNGGNG ARAARGGNGT NMGNATHAAR	1560
	GARTGYMGNG TNATGMGNGA YYTNAARGGN GTNCAYGGNA AYATGAARGG NCARWSNYTN	1620
	GGNTAYGCNT TYGCNGARTT YCARGARCAY GARCAYGCNY TNAARGCNYT NMGNYTNATH	1680
	AAYAAAYC CNGARATHTT YGGNCCNYTN AARMGNCCNA THGTNGARTT YWSNYTNGAR	1740
20	GAYMGNMGNA ARYTNAARAT GAARGARYTN MGNATHCARM GNWSNYTNCA RAARATGMGN	1800
	WSNAARCCNG CNACNGGNGA RCCNCARAAR GGNCARCCNG ARCCNGCNAA RGAYCARCAR	1860
	CARAARGCNG CNCARCAYCA YACNGARGAR CARWSNAARG TNCCNCCNGA RCARAARMGN	1920
	AARGCNGGNW SNACNWSNTG GACNGGNTTY CARACNAARG CNGARGTNGA RCARGTNGAR	1980
	YTNCNGAYG GNAARAARMG NMGNAARGTN YTNGCNYTNC CNWSNCAIMG NGGNCCNAAR	2040
25	ATHMGNYTNM GNGAYAARGG NAARGTNAAR CCNGTNCAYC CNAARAARCC NAARCCNCAR	2100
	ATHAAYCART GGAARCARGA RAARCRCAR YTNWSNWSNG ARCARGTNWS NMGNAARAAR	2160
	GCNAARGGNA AYAARACNGA RACNMGNTTY AAYCARYTNG TNGARCARTA YAARCARAAR	2220
	YTNYTNGGNC CNWSNAARGG NGCNCCNYTN GCNAARMGNW SNAARTGGTT YGAYWSN	2277

30 (2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2733 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

35 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

	TGAGCTTGGT	TGTCCTACCA	AAGCCAGCGT	TTCGGCTCGC	GTGCGCCGGC	CTAGTTTGCT	60
	CGCGTCCTCA	CGCGCTTTGG	GTTTCCCGGT	CTCATGGCCG	GCCTGACCTT	ATTTGTGGGC	120
	CGCCTCCCGC	CCTCGGCCCC	CAGTGAGCAG	CTGGAGGAAC	TGTTCAGTCA	GGTGGGGCCG	180
	GTGAAGCAGT	GCTTCGTGGT	GACTGAAAAA	GGGAGTAAGG	CATGTCGAGG	CTTTGGCTAT	240
	GTCACTTTTT	CAATGCTGGA	AGATGTTTCA	AGGGCCCTCA	AGGAGATTAC	CACCTTTGAA	300
5	GGTTGCAAGA	TCAACGTGAC	TGTTGCCAAG	AAAAAACTGA	GGAACAAGAC	AAAGGAAAAAG	360
	GGGAAAAATG	AAAACTCAGA	GTGCCCAAAG	AAGGAGCCGA	AGGCTAAAAA	AGCCAAAGTG	420
	GCAGATAAGA	AAGCCAGATT	AATTATTCGG	AACCTGAGCT	TTAAGTGTTT	AGAAGATGAC	480
	TTGAAGACAG	TATTTGCTCA	ATTTGGAGCT	GTCCTGGAAG	TAAATATCCC	TAGGAAACCA	540
	GATGGGAAGA	TGCGCGGTTT	TGGTTTTGTT	CAGTTCAAAA	ACCTCCTAGA	AGCAGGTAAA	600
10	GCTCTCAAAG	GCATGAACAT	GAAAGAGATA	AAAGGCCGGA	CAGTGGCTGT	GGATTGGGCC	660
	GTGGCAAAGG	ATAAATATAA	AGATACACAG	TCTGTTTCTG	CTATAGGTGA	GGAAAAGAGC	720
	CATGAATCTA	AACATCAGGA	ATCAGTTAAA	AAGAAGGGCA	GAGAGGAAGA	GGATATGGAA	780
	GAGGAAGAAA	ACGATGATGA	TGACGATGAT	GATGATGAAG	AAGATGGGGT	TTTTGATGAT	840
	GAAGATGAAG	AGGAAGAGAA	TATAGAATCA	AAGGTGACCA	AGCCTGTGCA	AATTCAGAAG	900
15	AGAGCAGTCA	AGAGACCAGC	CCCTGCAAAA	AGCAGTGATC	ATTCTGAGGA	GGACAGTGAC	960
	CTAGAGGAAA	GCGATAGTAT	TGATGATGGA	GAGGAACTGG	CTCAGAGTGA	TACCAGCACT	1020
	GAGGAGCAAG	AGGATAAAGC	TGTGCAAGTC	TCAAAACAAA	AGAAGAGGAA	ATTACCCTCT	1080
	GATGTGAATG	AAGGGAAAAC	TGTTTTTATC	AGAAATCTGT	CCTTTGACTC	AGAAGAAGAA	1140
	GAAGTTGGGG	AGCTTCTCCA	ACAGTTTGGA	GAAGTCAAAT	ATGTCCGCAT	TGTCTTGCAT	1200
20	CCAGACACAG	AGCATTCTAA	AGGTTGTGCA	TTTGCCCAAGT	TCATGACTCA	AGAAGCAGCT	1260
	CAGAAATGCC	TTCTAGCTGC	TTCTCCAGAG	AATGAGGCTG	GTGGGCTTAA	ACTGGATGGC	1320
	CGGCAGCTCA	AGGTTGACTT	GGCGGTGACC	CGTGATGAGG	CTGCAAAGCT	TCAGACGACG	1380
	AAGGTGAAGA	AGCCGACTGG	CACCCGGAAT	CTCTATCTGG	CCCGAGAAGG	CTTGATTTCGT	1440
	GCTGGGACGA	AGGCTGCAGA	GGGTGTGAGT	GCTGCTGATA	TGGCCAAAAG	AGAACGGTTT	1500
25	GAGCTGCTGA	AGCATCAGAA	ACTCAAGGAC	CAGAATATCT	TTGTCTCCCG	AACCAGGCTC	1560
	TGCCTGCACA	ATCTCCCAAA	GGCTGTAGAT	GACAAACAGC	TCAGAAAGCT	GCTGCTGAGT	1620
	GCTACTAGTG	GAGAGAAAAG	GGTGCGCATC	AAGGAGTGTA	GAGTGATGCG	AGACCTCAAA	1680
	GGAGTTCATG	GGAACATGAA	GGGTCAAGTC	CTGGGCTACG	CCTTTGCGGA	GTTCCAAGAG	1740
	CACGAGCATG	CCCTGAAAGC	CCTCCGCCTC	ATCAACAACA	ATCCAGAAAT	CTTTGGGCCT	1800
30	CTGAAGAGAC	CAATAGTGGA	GTTCTCTTTA	GAAGATCGAA	GAAAACTTAA	AATGAAGGAA	1860
	TTAAGGATCC	AGCGCAGCTT	GCAAAAAATG	AGATCCAAGC	CTGCAACTGG	TGAGCCTCAG	1920
	AAGGGGCAAC	CAGAGCCTGC	AAAAGACCAG	CAACAGAAGG	CAGCTCAACA	CCACACAGAG	1980
	GAACAAAGCA	AGGTGCCCCC	AGAGCAGAAG	AGAAAGGCGG	GCTCTACCTC	ATGGACCGGG	2040
	TTCCAGACCA	AGGCTGAAGT	GGAGCAGGTG	GAGCTGCCTG	ATGGAAAGAA	GAGAAGAAAG	2100
35	GTCCTGGCGC	TCCCCTCACA	CCGAGGCCCC	AAAATCAGGT	TGCGGGACAA	AGGCAAAGTG	2160
	AAGCCCGTCC	ATCCCAAAAA	GCCAAAGCCA	CAGATAAACC	AGTGGAAGCA	GGAGAAGCAG	2220
	CAATTATCGT	CCGAGCAGGT	ATCTAGGAAA	AAAGCTAAGG	GAAATAAGAC	GGAAACCCGC	2280

TTCAACCAGC TGGTCGAACA ATATAAGCAG AAATTATTGG GACCTTCTAA AGGAGCACCT 2340
 CTTGCAAAGA GGAGCAAATG GTTTGATAGT TGATGATGGC AGCAGGCTGG GTAAGAAGCT 2400
 GGGTTGTATA CTTTCTGGTG ACACTCCTGG GCTCCTCCCC ATCCCCCGTG TCTCTCACTG 2460
 AGGGAAAGAA AATCCCCAAG GGCAGTGCCA CTGTGCTCGG AGGTGCCCTG GACTGTGTAC 2520
 ATCTGAACTT TGGTCCATCC TTTGATGTGT GGTTCGTTAG CCACAAAGAG AAATATCTGA 2580
 5 AAGTCAACAT GATGCTTCTT GCATATTATC CAGATTATTG TATGAAGTTG TGTCTATAAT 2640
 TATTACCAAT TTTTATTCTT TATTTCTCAA ATGGAAACAC CTGAAAAAGC AAAAAAAAAA 2700
 AAAAAAAAAA CTCGAGGGGG GCCCGTACCC AAT 2733

(2) INFORMATION FOR SEQ ID NO:4:

10

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2277 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

15

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

AUGGCUGGUC UGACCCUGUU CGUUGGUCGU CUGCCGCCGU CCGCUCGUUC CGAACAGCUG 60
 GAAGAACUGU UCUCCCAGGU UGGUCCGGUU AAACAGUGCU UCGUUGUUAC CGAAAAAGGU 120
 UCCAAAGCUU GCCGUGGUUU CGGUUACGUU ACCUUCUCCA UGCUGGAAGA CGUUCAGCGU 180
 20 GCUCUGAAAG AAAUACCAC CUUCGAAGGU UGCAAAAUCA ACGUUACCGU UGCUAAAAAA 240
 AAACUGCGUA ACAAACCAA AGAAAAAGGU AAAAACGAAA ACUCCGAAUG CCCGAAAAAA 300
 GAACCGAAAG CUAAAAAGC UAAAGUUGCU GACAAAAAG CUCGUCUGAU CAUCCGUAAC 360
 CUGUCCUUA AAUGCUCGA AGACGACCUG AAAACCGUUU UCGCUCAGUU CGGUGCUGUU 420
 CUGGAAGUUA ACAUCCCGCG UAAACCGGAC GGUAAAAUGC GUGGUUUCGG UUUCGUUCAG 480
 25 UUCAAAAACC UGCUGGAAGC UGGUAAAGCU CUGAAAGGUA UGAACAUGAA AGAAAUCAAA 540
 GGUCGUACCG UUGCUGUUGA CUGGGCUGUU GCUAAAGACA AAUACAAAGA CACCCAGUCC 600
 GUUUCGCUA UCGGUGAAGA AAAUCCAC GAAUCCAAAC ACCAGGAAUC CGUUAAAAAA 660
 AAAGGUCGUG AAGAAGAAGA CAUGGAAGAA GAAGAAAACG ACGACGACGA CGACGACGAC 720
 GACGAAGAAG ACGGUGUUUU CGACGACGAA GACGAAGAAG AAGAAAACAU CGAAUCCAAA 780
 30 GUUACCAAAC CGGUUCAGAU CCAGAAACGU GCUGUUAAC GUCCGGCUCC GGCUAAAUCC 840
 UCCGACCACU CCGAAGAAGA CUCCGACCUG GAAGAAUCCG ACUCCAUCGA CGACGGUGAA 900
 GAACUGGCUC AGUCCGACAC CUCCACCGAA GAACAGGAAG ACAAAGCUGU UCAGGUUUC 960
 AACAAAAAAA AACGUAAACU GCCGUCCGAC GUUAACGAAG GUAAAACCGU UUUCAUCCGU 1020
 AACCUGUCCU UCGACUCCGA AGAAGAAGAA CUGGGUGAAC UGCUGCAGCA GUUCGGUGAA 1080
 35 CUGAAAUACG UUCGUUUCGU UCUGCACCCG GACACCGAAC ACUCCAAAGG UUGCGCUUUC 1140
 GCUCAGUUA UGACCCAGGA AGCUGCUCAG AAAUGCCUGC UGGCUGCUUC CCCGAAAAAC 1200
 GAAGCUGGUG GUCUGAAACU GGACGGUCGU CAGCUGAAAG UUGACCUGGC UGUUACCCGU 1260

	GACGAAGCUG	CUAAACUGCA	GACCACCAAA	GUUAAAAAAC	CGACCGGUAC	CCGUAACCUG	1320
	UACCUGGCUC	GUGAAGGUCU	GAUCCGUGCU	GGUACCAAAG	CUGCUGAAGG	UGUUUCCGCU	1380
	GCUGACAUGG	CUAAACGUGA	ACGUUUCGAA	CUGCUGAAAC	ACCAGAAACU	GAAAGACCAG	1440
	AACAUCUUCG	UUUCCCGUAC	CCGUCUGUGC	CUGCACAACC	UGCCGAAAGC	UGUUGACGAC	1500
	AAACAGCUGC	GUAAACUGCU	GCUGUCCGCU	ACCUCCGGUG	AAAAAGGUGU	UCGUAUCAAA	1560
5	GAAUGCCGUG	UUAUGCGUGA	CCUGAAAGGU	GUUCACGGUA	ACAUGAAAGG	UCAGUCCUG	1620
	GGUUACGCUU	UCCGUGAAUU	CCAGGAACAC	GAACACGCUC	UGAAAGCUCU	GCGUCUGAUC	1680
	AACAACAACC	CGGAAAUCUU	CGGUCCGUG	AAACGUCCGA	UCGUUGAAUU	CUCCUGGAA	1740
	GACCGUCGUA	AACUGAAAAU	GAAAGAACUG	CGUAUCCAGC	GUUCCUGCA	GAAAAUGCGU	1800
	UCCAAACCGG	CUACCGGUGA	ACCGCAGAAA	GGUCAGCCGG	AACCGGCUAA	AGACCAGCAG	1860
10	CAGAAAGCUG	CUCAGCACCA	CACCGAAGAA	CAGUCCAAAG	UUCCGCCCGA	ACAGAAACGU	1920
	AAAGCUGGUU	CCACCUCUG	GACCGGUUUC	CAGACCAAAG	CUGAAGUUGA	ACAGGUUGAA	1980
	CUGCCGGACG	GUAAAAAAG	UCGUAAAGUU	CUGGCUCUGC	CGUCCACCG	UGGUCCGAAA	2040
	AUCCGUCUGC	GUGACAAAGG	UAAAGUUAAA	CCGGUUCACC	CGAAAAAAC	GAAACCGCAG	2100
	AUCAACCAGU	GGAAACAGGA	AAAACAGCAG	CUGUCCUCCG	AACAGGUUUC	CCGUAAAAAA	2160
15	GCUAAAGGUA	ACAAAACCGA	AACCCGUUUC	AACCAGCUGG	UGAACAGUA	CAAACAGAAA	2220
	CUGCUGGGUC	CGUCCAAAGG	UGCUCGCGUG	GCUAAACGUU	CCAAUUGGUU	CGACUCC	2277

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- 20 (A) LENGTH: 2277 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

	ATGGCCGGCC	TGACCCTGTT	CGTGGGCCGC	CTGCCCCCA	GCGCCCGCAG	CGAGCAGCTG	60
	GAGGAGCTGT	TCAGCCAGGT	GGGCCCCGTG	AAGCAGTGCT	TCGTGGTGAC	CGAGAAGGGC	120
	AGCAAGGCCT	GCCGCGGCTT	CGGCTACGTG	ACCTTCAGCA	TGCTGGAGGA	CGTGCAGCGC	180
	GCCCTGAAGG	AGATCACCAC	CTTCGAGGGC	TGCAAGATCA	ACGTGACCGT	GGCCAAGAAG	240
30	AAGCTGCGCA	ACAAGACCAA	GGAGAAGGGC	AAGAACGAGA	ACAGCGAGTG	CCCCAAGAAG	300
	GAGCCCAAGG	CCAAGAAGGC	CAAGGTGGCC	GACAAGAAGG	CCCGCCTGAT	CATCCGCAAC	360
	CTGAGCTTCA	AGTGCAGCGA	GGACGACCTG	AAGACCGTGT	TCGCCCAGTT	CGGCGCCGTG	420
	CTGGAGGTGA	ACATCCCCCG	CAAGCCCGAC	GGCAAGATGC	GCGGCTTCGG	CTTCGTGCAG	480
	TTCAAGAACC	TGCTGGAGGC	CGGCAAGGCC	CTGAAGGGCA	TGAACATGAA	GGAGATCAAG	540
35	GGCCGCACCG	TGGCCGTGGA	CTGGGCCGTG	GCCAAGGACA	AGTACAAGGA	CACCCAGAGC	600
	GTGAGCGCCA	TCGGCGAGGA	GAAGAGCCAC	GAGAGCAAGC	ACCAGGAGAG	CGTGAAGAAG	660
	AAGGGCCGCG	AGGAGGAGGA	CATGGAGGAG	GAGGAGAACG	ACGACGACGA	CGACGACGAC	720

	GACGAGGAGG ACGGCGTGTT CGACGACGAG GACGAGGAGG AGGAGAACAT CGAGAGCAAG	780
	GTGACCAAGC CCGTGCAGAT CCAGAAGCGC GCCGTGAAGC GCGCCGCCCC CGCCAAGAGC	840
	AGCGACCACA GCGAGGAGGA CAGCGACCTG GAGGAGAGCG ACAGCATCGA CGACGGCGAG	900
	GAGCTGGCCC AGAGCGACAC CAGCACCGAG GAGCAGGAGG ACAAGGCCGT GCAGGTGAGC	960
	AACAAGAAGA AGCGCAAGCT GCCCAGCGAC GTGAACGAGG GCAAGACCGT GTTCATCCGC	1020
5	AACCTGAGCT TCGACAGCGA GGAGGAGGAG CTGGGCGAGC TGCTGCAGCA GTTCGGCGAG	1080
	CTGAAGTACG TGCGCATCGT GCTGCACCCC GACACCGAGC ACAGCAAGGG CTGCGCCTTC	1140
	GCCCAGTTCA TGACCCAGGA GGCCGCCAG AAGTGCCTGC TGGCCGCCAG CCCCAGAAAC	1200
	GAGGCCGGCG GCCTGAAGCT GGACGGCCGC CAGCTGAAGG TGGACCTGGC CGTGACCCGC	1260
	GACGAGGCCG CCAAGCTGCA GACCACCAAG GTGAAGAAGC CCACCGGCAC CCGCAACCTG	1320
10	TACCTGGCCC GCGAGGGCCT GATCCGCGCC GGCACCAAGG CCGCCGAGGG CGTGAGCGCC	1380
	GCCGACATGG CCAAGCGCGA GCGCTTCGAG CTGCTGAAGC ACCAGAAGCT GAAGGACCAG	1440
	AACATCTTCG TGAGCCGCAC CCGCCTGTGC CTGCACAACC TGCCCAAGGC CGTGGACGAC	1500
	AAGCAGCTGC GCAAGCTGCT GCTGAGCGCC ACCAGCGGCG AGAAGGGCGT GCGCATCAAG	1560
	GAGTGCCGCG TGATGCGCGA CCTGAAGGGC GTGCACGGCA ACATGAAGGG CCAGAGCCTG	1620
15	GGCTACGCCT TCGCCGAGTT CCAGGAGCAC GAGCAGCCCC TGAAGGCCCT GCGCCTGATC	1680
	AACAACAACC CCGAGATCTT CGGCCCCCTG AAGCGCCCCA TCGTGGAGTT CAGCCTGGAG	1740
	GACCGCCGCA AGCTGAAGAT GAAGGAGCTG CGCATCCAGC GCAGCCTGCA GAAGATGCGC	1800
	AGCAAGCCCC CCACCGCGCA GCCCCAGAAG GGCCAGCCCC AGCCCCGCAA GGACCAGCAG	1860
	CAGAAGGCCG CCCAGCACCA CACCGAGGAG CAGAGCAAGG TGCCCCCGA GCAGAAGCGC	1920
20	AAGGCCGGCA GCACCAGCTG GACCGGCTTC CAGACCAAGG CCGAGGTGGA GCAGGTGGAG	1980
	CTGCCCCAGC GCAAGAAGCG CCGCAAGGTG CTGGCCCTGC CCAGCCACCG CGGCCCCAAG	2040
	ATCCGCCTGC GCGACAAGGG CAAGGTGAAG CCCGTGCACC CCAAGAAGCC CAAGCCCCAG	2100
	ATCAACCACT GGAAGCAGGA GAAGCAGCAG CTGAGCAGCG AGCAGGTGAG CCGCAAGAAG	2160
	GCCAAGGGCA ACAAGACCGA GACCCGCTTC AACCAGCTGG TGGAGCAGTA CAAGCAGAAG	2220
25	CTGCTGGGCC CCAGCAAGGG CGCCCCCTG GCCAAGCGCA GCAAGTGGTT CGACAGC	2277

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 540 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

35	GGGTTGCGGA GGGTGGGCCT GGGAGGGGTG GTGGCCATTT TTTGTCTAAC CCTAACTGAG	60
	AAGGGCGTAG GCGCCGTGCT TTTGCTCCCC GCGCGCTGTT TTTCTCGCTG ACTTTCAGCG	120
	GGCGGAAAAG CCTCGGCCTG CCGCCTTCCA CCGTTCATTC TAGAGCAAAC AAAAAATGTC	180

AGCTGCTGGC CCGTTCGCCC CTCCCGGGGA CCTGCGGCGG GTCGCCTGCC CAGCCCCCGA 240
 ACCCCGCCTG GAGGCCGCGG TCGGCCCGGG GCTTCTCCGG AGGCACCCAC TGCCACCGCG 300
 AAGAGTTGGG CTCTGTCAGC CGCGGGTCTC TCGGGGGCGA GGGCGAGGTT CAGGCCTTTC 360
 AGGCCGAGG AAGAGGAACG GAGCGAGTCC CCGCGCGCGG CGCGATTCCC TGAGCTGTGG 420
 GACGTGCACC CAGGACTCGG CTCACACATG CAGTTCGCTT TCCTGTTGGT GGGGGGAACG 480
 5 CCGATCGTGC GCATCCGTCA CCCCTCGCCG GCAGTGGGGG CTTGTGAACC CCCAAACCTG 540

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 540 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

15 GGGTTGCCGA GGGTGGGCCT GGGAGGGGTG GTGGCCATTT TTTGTCCAAC CCCAACTGAG 60
 AAGGGCGTAG GCGCCGTGCT TTTGCTCCCC GCGCGCTGTT TTTCTCGCTG ACTTTCAGCG 120
 GGCGGAAAAG CCTCGGCCTG CCGCCTTCCA CCGTTCATTC TAGAGCAAAC AAAAAATGTC 180
 AGCTGCTGGC CCGTTCGCCC CTCCCGGGGA CCTGCGGCGG GTCGCCTGCC CAGCCCCCGA 240
 ACCCCGCCTG GAGGCCGCGG TCGGCCCGGG GCTTCTCCGG AGGCACCCAC TGCCACCGCG 300
 20 AAGAGTTGGG CTCTGTCAGC CGCGGGTCTC TCGGGGGCGA GGGCGAGGTT CAGGCCTTTC 360
 AGGCCGAGG AAGAGGAACG GAGCGAGTCC CCGCGCGCGG CGCGATTCCC TGAGCTGTGG 420
 GACGTGCACC CAGGACTCGG CTCACACATG CAGTTCGCTT TCCTGTTGGT GGGGGGAACG 480
 CCGATCGTGC GCATCCGTCA CCCCTCGCCG GCAGTGGGGG CTTGTGAACC CCCAAACCTG 540

25 (2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- 30 (A) LENGTH: 540 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GGGTTGCCGA GGGTGGGCCT GGGAGGGGTG GTGGCCATTT TTTGTCTAAG CCTAAGTGAG 60
 AAGGGCGTAG GCGCCGTGCT TTTGCTCCCC GCGCGCTGTT TTTCTCGCTG ACTTTCAGCG 120
 35 GGCGGAAAAG CCTCGGCCTG CCGCCTTCCA CCGTTCATTC TAGAGCAAAC AAAAAATGTC 180
 AGCTGCTGGC CCGTTCGCCC CTCCCGGGGA CCTGCGGCGG GTCGCCTGCC CAGCCCCCGA 240
 ACCCCGCCTG GAGGCCGCGG TCGGCCCGGG GCTTCTCCGG AGGCACCCAC TGCCACCGCG 300

AAGAGTTGGG CTCTGTCAGC CGCGGGTCTC TCGGGGGCGA GGGCGAGGTT CAGGCCTTTC 360
 AGGCCGAGG AAGAGGAACG GAGCGAGTCC CCGCGCGCGG CGCGATTCCC TGAGCTGTGG 420
 GACGTGCACC CAGGACTCGG CTCACACATG CAGTTCGCTT TCCTGTTGGT GGGGGGAACG 480
 CCGATCGTGC GCATCCGTCA CCCCTCGCCG GCAGTGGGGG CTTGTGAACC CCCAAACCTG 540

5 (2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 538 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 10 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GGGTTGCGGA GGGTGGGCCT GGGAGGGGTG GTGGCCATTT TTTGTCTACC CTA CTGAGAA 60
 GGGCGTAGGC GCCGTGCTTT TGCTCCCCGC GCGCTGTTTT TCTCGCTGAC TTTCAGCGGG 120
 15 CGGAAAAGCC TCGGCCTGCC GCCTTCCACC GTTCATTCTA GAGCAAACAA AAAATGTCAG 180
 CTGCTGGCCC GTTCGCCCCCT CCCGGGGACC TCGGGCGGGT CGCCTGCCCA GCCCCGAAC 240
 CCCGCCTGGA GGCCGCGGTC GGCCCGGGGC TTCTCCGGAG GCACCCACTG CCACCCGCGAA 300
 GAGTTGGGCT CTGTCAGCCG CGGGTCTCTC GGGGGCGAGG GCGAGGTTCA GGCCTTTTCAG 360
 GCCGCAGGAA GAGGAACGGA GCGAGTCCCC GCGCGCGGCG CGATTCCCTG AGCTGTGGGA 420
 20 CGTGACCCCA GGA CTGCGCT CACACATGCA GTTCGCTTTC CTGTTGGTGG GGGGAACGCC 480
 GATCGTGCGC ATCCGTACCC CCTCGCCGGC AGTGGGGGCT TGTGAACCCC CAAACCTG 538

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- 25 (A) LENGTH: 14 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: not relevant
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

30 (ix) FEATURE:

- (A) NAME/KEY: Peptide
 (B) LOCATION: 5..13
 (D) OTHER INFORMATION: /note= "Xaa represents isoleucine
 or leucine"

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Glu Ala Ala Thr Xaa Asp Xaa Pro Gln Gln Gly Ala Asn Lys

1

5

10

WHAT IS CLAIMED IS:

1. An isolated protein comprising a telomerase protein selected from the group consisting of human telomerase protein p140, p105, p48 and p43, having molecular weights of about 140kD, about 105kD, about 48kD and about 43kD, respectively, as determined by polyacrylamide gel electrophoresis under denaturing conditions, or a human telomerase protein domain thereof having telomerase-specific activity.
2. An isolated protein according to claim 1, wherein said protein specifically binds at least one of the telomerase RNA of SEQ ID NO:6, a telomerase primer, or a nucleotide triphosphate.
3. An isolated protein according to claim 1, said protein isolated by:
lysing HeLa cells by dounce homogenization in low-salt buffer to produce cell lysates and
centrifuging said lysates supplemented with 15% glycerol at 125,000 x g for 50 minutes to obtain a first soluble fraction enriched for telomerase (S-100 fraction);
binding said S-100 fraction adjusted to 0.2 M ammonium sulfate to SP Sepharose (Pharmacia), and developing with a gradient in sodium chloride, to obtain a second soluble fraction enriched for telomerase (SP fraction);
binding said SP fraction adjusted to about 0.3-0.4 M ionic strength to Q Sepharose (Pharmacia) adjusting said SP fraction, and developing with a gradient in sodium chloride, to obtain a third soluble fraction enriched for telomerase (Q fraction); binding said Q fraction adjusted to about 0.3-0.4 M ionic strength to phosphocellulose (Whatman), and developing in sodium chloride, to obtain a fourth soluble fraction enriched for telomerase (PC fraction);
binding said PC fraction adjusted to about 0.3-0.4 M ionic strength to 2'Omethyl RNA oligonucleotide immobilized on streptavidin agarose (Sigma) and eluting with a electrophoresis sample medium comprising 5% β -mercaptoethanol and 2% Sodium Dodecyl Sulfate to obtain a fifth soluble fraction (2'Omethyl fraction);
separating said 2'Omethyl fraction by polyacrylamide gel electrophoresis under denaturing conditions to obtain a gel band at a molecular weight of about 140kD, 105kD,

48kD or 43kD having telomerase activity;

excising or eluting said gel band to obtain a human telomerase protein of a molecular weight of about 140kD, 105kD, 48kD or 43kD, respectively, as determined by polyacrylamide gel electrophoresis under denaturing conditions.

5 4. An isolated protein comprising a portion of the amino acid sequence of SEQ ID NO:1 sufficient for telomerase-specific activity.

10 5. An isolated protein according to claim 4, wherein said protein comprises a human telomerase protein domain selected from the group consisting of an RNA binding domain, a telomerase subunit binding domain, and a substrate, agonist, antagonist, chaperone or cytoskeletal binding domain.

6. A recombinant nucleic acid encoding a protein according to claim 1 or 4.

15 7. A cell comprising a nucleic acid according to claim 6.

20 8. A method of making an isolated telomerase protein, comprising steps: introducing a nucleic acid according to claim 6 into a host cell or cellular extract, incubating said host cell or extract under conditions whereby said nucleic acid is expressed as a transcript and said transcript is expressed as a translation product comprising said protein, and isolating said translation product.

9. An isolated human telomerase protein made by the method of claim 8.

25 10. An isolated human telomerase nucleic acid comprising SEQ ID NO:6, or a fragment thereof sufficient to specifically hybridize with, or amplify from a nucleic acid having the sequence defined by SEQ ID NO:6.

30 11. An isolated telomerase nucleic acid according to claim 10 comprising at least one of SEQ ID NO:6, bases 191-210, bases 245-259, bases 341-369 and bases 381-399.

12. A method of screening for an agent which modulates the binding of a human telomerase protein to a binding target, said method comprising the steps of:

incubating a mixture comprising:

an isolated protein according to claim 1,

a binding target of said protein, and

5 a candidate agent;

under conditions whereby, but for the presence of said agent, said protein specifically binds said binding target at a reference affinity;

detecting the binding affinity of said protein to said binding target to determine an agent-biased affinity,

10 wherein a difference between the agent-biased affinity and the reference affinity indicates that said agent modulates the binding of said protein to said binding target.

13. A method according to claim 12, wherein said binding target is a substrate of said protein and said reference and agent-biased binding affinity are each detected as the
15 polymerization by said protein of a nucleic acid on said substrate.

14. An isolated somatically recombined protein receptor which specifically binds a protein according to claim 1, wherein said receptor is an antibody or a T-cell antigen
20 receptor.

15. A method of modulating the binding of a human telomerase or telomerase protein to a human telomerase binding target, said method comprising contacting said telomerase or telomerase protein with a receptor according to claim 14.

25 16. A method according to claim 15, wherein said binding target is a substrate of said telomerase and said receptor inhibits the polymerization by said telomerase of a nucleic acid on said substrate.

17. A method of polymerizing a nucleic acid on a substrate, comprising contacting said
30 substrate with a telomerase comprising a protein according to claim 1, wherein said telomerase comprises a nucleic acid template having a preselected nucleotide sequence.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US97/12296

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : Please See Extra Sheet.

US CL : Please See Extra Sheet.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/194, 240.1, 252.3, 320.1, 69.1, 91.3, 172.3, 7.1; 530/350; 536/23.1, 23.2, 24.31, 24.33

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y,P	US 5,583,016 A (VILLEPONTEAU et al.) 10 December 1996, entire patent, especially the abstract and column 20, lines 10-60.	1-17
Y	WO 96/19580 A2 (COLD SPRING HARBOR LABORATORY) 27 June 1996. See abstract and examples 7-10.	1-17
Y	COUNTER et al. Telomerase Activity in Human Ovarian Carcinoma. Proc. Natl. Acad. Sci. USA. April 1994. Vol. 91, pages 2900-2904, see entire article.	1-17

☐ Further documents are listed in the continuation of Box C.
 ☐ See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
A document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
E earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*A* document member of the same patent family
O document referring to an oral disclosure, use, exhibition or other means	
P document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

04 SEPTEMBER 1997

Date of mailing of the international search report

28 OCT 1997

 Name and mailing address of the ISA/US
 Commissioner of Patents and Trademarks
 Box PCT
 Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

TEKCHAND SAIDHA

Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US97/12296

A. CLASSIFICATION OF SUBJECT MATTER:

IPC (6):

C12N 9/12, 5/00, 1/20, 15/00; C12P 21/06, 19/34; C07K 1/00; C07H 21/02, 21/04

A. CLASSIFICATION OF SUBJECT MATTER:

US CL :

435/194, 240.1, 252.3, 320.1, 69.1, 91.3, 172.3, 7.1; 530/350; 536/23.1, 23.2, 24.31, 24.33

B. FIELDS SEARCHED

Electronic data bases consulted (Name of data base and where practicable terms used):

APS, STN Files : Medline, Caplus, Biosis, Wpids, Biotechds, Scisearch & Biotechds. Search terms : Telomerase and (DNA or RNA or protein), and human, telomerase, etc. Protein and Nucleic Acid data base search for the amino acid and DNA sequences.